The notion of data worlds is thus partly a response to contemporary "socio-technical imaginaries" (Jasanoff and Kim 2015) about data. Just as industrial technologies of the past were accompanied by new social, cultural and political imaginaries, so we can trace the ascent of "data imaginaries" and "data speak": visions and rhetoric concerning the role of data in society. As Gillespie notes in relation to platforms, these imaginaries do "discursive work" (2010). For example, data is framed as "the new oil", "the new gold" or "the new soil", in order to emphasise its value as a social or economic *resource*. We also see the idea of "infrastructures" of data being used in order to emphasise different configurations of public-private and state-citizen collaboration, as well as to establish information infrastructures as a basic good in society alongside infrastructures for water, gas, electricity and so on. The platform, the portal, the app, the lab and the hackday give rise to new imaginaries and discursive regimes as well as material practices suggesting the role of data in public life.

Many of these imaginaries focus on the *value* to be extracted from data, through various mechanisms and arrangements to *make data public*. The issue is often framed as one of access, formats and conventions for encouraging the re-use of public data in innovative applications and services. I have found the notion of data worlds to be useful in examining what open data initiatives do and do not do, and how they might be done differently (Gray 2018). For example, open data projects may focus on *redistributing access* to data about public finances without substantively engaging with the epistemic, social and political work of data infrastructures in selecting, translating, arranging and articulating certain aspects of fiscal policy (such as detailed spending estimates of local councils), but not others (such as the economic activities and tax payments of multinational corporations).

The notion of data worlds is intended to gesture beyond two prominent forms of data politics which can be broadly characterised as "data liberation" and "data protection". Both emphasise dynamics of power related to *access*. Data liberation is widely associated with hacker culture and other forms of information activism: setting information free from institutions and corporations as a means to address information asymmetries, and to provide a resource for activism and social change. This may be considered in terms of a "Promethean" mythology of broadening access to a powerful resource or instrument (just as Prometheus stole fire from the gods to give to humankind) – an outlook which is shared across the spectrum from the "mega-leaks" of Wikileaks and the Panama Papers and the more curated, selective leaks of Edward Snowden, through to "Freedom of Information" (FOI) and access to information movements in the 1990s, as well as official and grassroots open data initiatives which emphasise making data legally and technically amenable to re-use (Gray 2016).

On the flipside, we have "data protection" as a narrative of information politics which emphasises the protection of personal information from state, corporate and other actors – as is exemplified in the work of civil society groups such as Privacy International and the Electronic Frontier Foundation. These narratives place an emphasis on the individual ownership and control of personal data, as well as on preventing, obstructing, managing, regulating and raising awareness of the collection of personally identifying information – from artistic projects to make visible the personal information different actors have collected, to law and policy (such as EU Data Protection rules or US Fair Information Principles).

Whilst these two genres of information policy and information politics are indeed vital, data infrastructures do much more than *making data public* and *making data private*. Raymond Geuss has critiqued what he considers the disproportionate attention accorded to the "public/private distinction" which both reflects and reinforces the absence of "any effective general framework for thinking about politics apart from liberalism" (Geuss 2003). Dominic Boyer has suggested the phrase "digital liberalism" as an invitation to attend to how "techno-institutional processes such as computerization and digital information and politico-institutional discourses of late liberalism have coevolved, at times reinforcing and naturalizing each other, promoting novel bundles of epistemics and ethics" (Boyer 2013).

The notion of data worlds is intended to make space for thinking about data as more than simply a representational resource, and the politics of data as more than a matter of liberation and protection. It is intended to encourage exploration of

the *performative* capacities of data infrastructures: what they do and could do differently, and how they are done and could be done differently. This includes consideration of, as Geoffrey Bowker puts it, "the ways in which our social, cultural and political values are braided into the wires, coded into the applications and built into the databases which are so much a part of our daily lives" (2014). In doing so we may draw on performative analyses of numbers (Espeland and Stevens 2008; Verran 2015), models (Mackenzie 2008) and methods (Law, Ruppert and Savage 2011) to consider how data infrastructures may be involved in not just the *representation* but also the *articulation* of collective life, while at the same time being the products of social and institutional work themselves.

Many accounts of performativity allude to the work of J. L. Austin, who suggests "the issuing of an utterance is the performing of an action" (1975: 6). Austin is associated with a "linguistic turn" in Anglophone analytic philosophy said to begin with Wittgenstein, whose later work reflects on what language does beyond referring to things. In the following discussion of the performative and world-making capacities of data infrastructures, I draw on an earlier linguistic turn that occurred in German philosophy in the eighteenth century and which has recently begun to receive more attention in English-language scholarship (Lifschitz 2012; Bowie 2013; Taylor 2016). Thinkers associated with this earlier turn also sought to look beyond representational accounts of language towards its other capacities as a situated set of social practices. Ian Hacking argues this tradition can be viewed as an alternative to Wittgenstein's "depoliticized" philosophy of language (2002). I do not argue for the special relevance of this period and these ideas. Rather I suggest that it contains conceptual and theoretical resources which may be useful when considering different aspects of worlds, worlding and world-making in relation to data.

The three aspects of data worlds which I examine below are not intended to be comprehensive, but illustrative of what is involved in data infrastructures, what they do, and how they are put to work. As I shall return to in the conclusion, this outline is intended to open up space for not only *thinking about data differently*, but also *doing things with data differently*. The test of these three aspects is therefore not only their analytical purchase, but also their practical utility.

1. Data Worlds as Horizons of Intelligibility

The first aspect of data worlds draws on philosophical ideas about worlds, worlding and world-making to look at how things are *sayable, knowable, intelligible* and *experiencable* through data. In European philosophy this begins with Kant's "Copernican revolution" which recognises the active and creative role that human beings played in composing the worlds that they experience – including through schemes, categories and structures such as space, time, causality and quantity which give form to experience. This is an explicit departure from views which saw experience as "given" and immediate, and also heralds a broader philosophical shift towards looking at how experience is articulated and mediated through language, culture and social arrangements.

Subsequent thinkers in this tradition –Hamann and Herder in the eighteenth century to thinkers as diverse as Heidegger, Gadamer, Benjamin and Wittgenstein in the twentieth century – stripped Kant's project of its aspiration to clarify universal structures, and highlighted the role of socially and historically situated linguistic and cultural infrastructures, or what the contemporary philosopher Charles Taylor calls "meaningful media" (Taylor 1985), in shaping our apprehension of the worlds we inhabit.

Many of these earlier thinkers mainly focused on the role of *language* as a horizon of intelligibility, providing the "conditions of possibility" for our experience. As Taylor notes, this also corresponded with an explicit move away from a dominant focus on the designative, representational and "information encoding" capacities of language and other meaningful media – and a focus on their role in composing and co-producing our worlds of experience (2016). As Hacking puts it, in this tradition we find the notion that: "language is creative; to it we owe the existences and structures that populate our world-versions" (2002, 139). And yet, while there is a

focus on language as an important and paradigmatic case of how our experience is formed, language is very often construed in a broad sense – including not only written and verbal language, but also music, painting, sculpture, and other social and cultural conventions for making meaning.

Benjamin draws on Hamann's "metacritical" challenge towards narrower conceptions of experience as "naked, primitive, self-evident" (Benjamin 1996), exploring in his work the world-making capacities of architecture, urban planning, fashion, advertising and technologies, perhaps most famously in his *Arcades Project* (Benjamin 1999). Later in the twentieth century, these kinds of appropriations of Kantian ideas about schematism and world-making (minus the transcendental idealist baggage), have broadened out from what Apel calls the "linguistic *a priori*" of thinkers like Hamann, Herder and Heidegger (Apel 1973, 39), to the "historical *a priori*" of Michel Foucault (Foucault 1972) and what has been called the "technological *a priori*" of German media theorists shifting the focus to *Kulturtechniken* or "cultural techniques" (Tuschling 2016; Winthrop-Young, Iurascu and Parikka 2013).

What might this sense of world-making bring to an understanding of the politics of data? Taking a cue from this theoretical constellation, we might envisage data worlding in terms of a contingent, historically and socially situated, technologically mediated "data *a priori*" which not only designates but also provides the conditions of possibility for seeing and engaging with different aspects of collective life – making possible particular styles of reasoning and particular forms of knowledge and experience.

Data practices might be understood not just in terms of more sophisticated ways of *picking things out*, but as contributing to new ways of *making things up*, as Hacking puts it (1985). Here critical data scholars can benefit from decades of research on social practices of quantification (Porter 1996; Espeland and Stevens 2008; Rottenburg, Merry, Park and Mugler 2015; Bruno, Jany-Catrice and Touchelay 2016); statistics (Porter 1986; Desrosières 2002); standards (Lampland and Star 2008); probability (Hacking 1990); visual reasoning (Halpern 2015); and other studies of cultures and practices of knowledge which focus not just on what is said, but on the background against which things become sayable. In looking at how data worlds provide horizons of intelligibility we can both draw on genealogies of the modes of experience and styles of reasoning which are rendered possible through data over previous decades and centuries, as well as looking at what is distinctive about new and emerging digital technologies. As Nelson Goodman puts it in his classic *Ways of Worldmaking*: "worldmaking as we know it always starts from worlds already on hand; the making is a remaking" (Goodman 1978, 6).

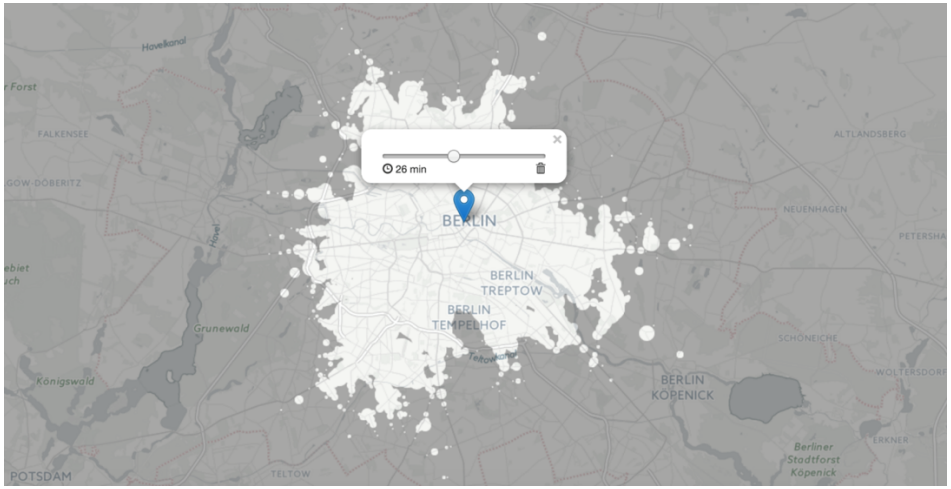

Figure 1: Screenshot of transportation maps from Mapnificent project (mapnificent.net), showing which places are accessible in a given amount of time from a given point.

Thus in relation to digital data worlds we may examine how composites of conventions, norms, technologies, practices, methods, pieces of software, graphical user interfaces, data standards, data formats and aesthetic approaches are implicated in making things up and making things intelligible with data. This might include looking at how horizons of intelligibility change from pre-digital to digital data worlds. For example, we might look at differences in how the world is conceptually organised or "carved up" into categories. In contrast to the classificatory practices of statisticians taking measure of economies or populations, "born digital" and big data, generated as a result of interactions with online platforms, can give rise to

novel practices of semi-automated classification, as well as emerging forms of inequality and discrimination.

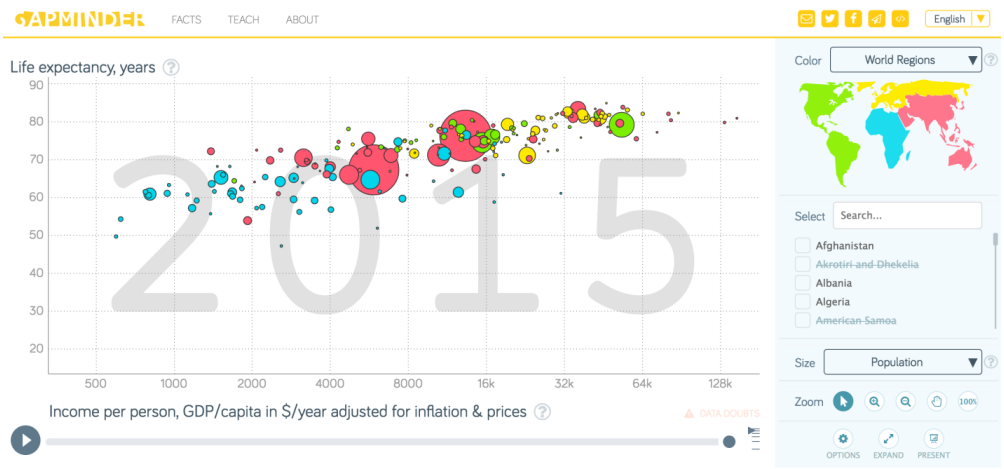

Figure 2: Screenshot of interactive "animated bubble charts" of Gapminder project (gapminder.org), exploring relations between average life expectancy and income per capita over time for countries around the world.

There are many historical studies looking at how social categories are articulated through statistical practices (Desrosières 2002). In digital data worlds computational techniques such as machine learning may be used to facilitate "class discovery". Clusters and orderings of hashtags, links, likes, images and other media can be viewed as *co-produced* by the logic of platforms, algorithms, and the "device cultures" of users. For example, Sam Lavigne's "Taxonomy of Humans According to Twitter" at *The New Inquiry* explores and visually represents the "bizarre rubrics Twitter uses to render its users legible" (Lavigne 2017). This project aims to make visible the way in which people are classified according to a combination of user activity and information from data-brokerage companies, leading to categories such as "people who live with three other people", "buyers of frozen ethnic foods", and "households whose behavior indicates they are spa mavens". These algorithmically-mediated data practices around online platforms can be understood, as Annemarie Mol puts it, as "new ways of *doing* reality" (Mol 1999).

We might also look at the forms of experience, styles of reasoning, and genres of sociality that arise with novel kinds of cultural objects associated with digital data worlds. This includes the world-making capacities of things such as apps, platforms, software packages, code libraries, and data analysis and visualisation tools through which people make sense with data, and integrate data into different kinds of social processes, practices and institutions. We might consider how space, time, relations and categories are articulated and organised through lists, tables, charts, timelines, maps and coordinate systems – and inscribed into dashboards, interactive data visualisations, word clouds, network graphs, mapping technologies, and computational techniques for filtering, reconciling and analysing data.

Just as Scott talks of "seeing like a state" by reducing "an infinite array of detail to a set of categories that will facilitate summary descriptions, comparisons, and aggregation" (1999, 77), and Law talks of "seeing like a survey" by using statistical methods to enact "a very particular version of the collective" (2009), so we may consider how the performative and world-making capacities of data projects are conventionalised into familiar forms such as seeing like an app, a network graph, a data portal, an API, an interactive map, a Google Spreadsheet and so on.

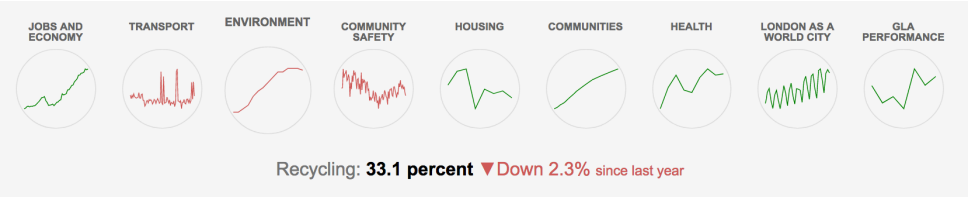

Figure 3: Detail of dashboard previews from London Datastore (data.london.gov.uk) showing trends in relation to performance indicators for the city.

"Time travel" maps articulate novel and interactive relationships between space and time by estimating the zones that can be reached from a given point in a given time interval (Figure 1). Global indicators are no longer simply represented through tables, charts or line graphs, but through interactive animated graphics dramatizing the passage of centuries through the movements of multicoloured bubbles articulating different dimensions of collective life (Figure 2). Interactive dashboards are

envisaged as the preferred mode to increase transparency and public accountability in the city by tracking trends in relation to key performance indicators (Figure 3). While these kinds of projects draw on ideals and practices that have much longer histories – such as the aspiration for what Theodore Porter characterises as "thin descriptions" and an aesthetic of distance – digital technologies are also facilitating reconfigurations and redistributions of these data world-making capacities, leading to emerging genres of making sense with data. As we shall see in the following section, these meaning-making practices should be understood as social conventions.

2. Data Worlds as Collective Accomplishments

The second aspect of data worlds draws on a sociological tradition of studying "social worlds". Adele Clarke and Susan Leigh Star trace this from the Chicago School of Sociology to Science and Technology Studies (Clarke and Star 2008). This approach encompasses and informs a range of research on social worlds – including for example Anselm Strauss, who suggests in the 1970s that we should look at the "social worlds" of genetics, high energy physics, computerisation and banking (Strauss 1978), to Howard Becker's renowned work on "art worlds" (Becker 1984), as well as the "worlds of classification" and "information worlds" explored in the work of Bowker, Star and other scholars of information infrastructures (Bowker and Star 2000; Star, Bowker and Neumann 1997).

This view of social world-making is also commensurate with both critics and radical interpreters of Kant who suggest that language and meaning-making practices should be regarded in fundamentally social and historical terms – a move which led Ian Hacking to mark this as a key moment when language "goes public" (2002). This tendency to look at language and meaning-making practices in terms of contingent and evolving social institutions is also present in Wittgenstein's work, which is a formative influence on subsequent social research agendas from ethnomethodology, to the "Strong Programme", to Science and Technology Studies (see, e.g. Bloor 1983, 2002; Hacking 1984; Lynch 1992).

Taking a cue from this tradition, we might look at how the information products, styles of reasoning, and meaning-making capacities associated with data infrastructures can be considered as "relational achievements" or "distributed accomplishments" – and how the collectives associated with data infrastructures are changing in composition. Data worlds as horizons of intelligibility must thus be understood as social and collective. Changes in these collectives can carry significant political and political-economic consequences. For example, in the case of the redistribution of "data work" from official institutions to actors outside the public sector – as in the case of open data initiatives (Gray 2018), to civil society and citizen generated data (Gray, Lämmerhirt and Bounegru, 2016), through a shift of emphasis from statistical data to "big data" generated by major technology companies (Flyverbom, Madsen and Rasche 2017).

In Howard Becker's terms, we can examine the "conventions" and practices which hold these social "data worlds" together – which he characterises as "ways of seeing and hearing that were known by everyone involved and thus formed the basis for their collective action" (Becker 1984, xv). In the case of open data, this might include, for example, such things as open licensing practices, legal arrangements, and technical practices which aim to "unlock the potential" of data as a resource, and "reduce the barriers" to its re-use by non-state actors – whether in new technology products such as Google Maps, the stories of data journalists, or the campaigns of NGOs or civil society groups. This concern with legal and technical conventions suggests that the open data community might be understood as what Chris Kelty calls a "recursive public", or "a public that is vitally concerned with the material and practical maintenance and modification of the technical, legal, practical, and conceptual means of its own existence as a public" (Kelty 2005, 3). There are also emerging conventions for making sense with data such as those discussed in the previous section.

Looking at data worlds as collective accomplishments includes recognising the role of actors whose contributions may otherwise be under-recognised. In his work on the sociology of "art worlds" Becker suggests a shift of emphasis from the formal quality of art works to "complex networks through which art happens" (1984, 1).

In his work he describes a broad range of materials, formats, spaces, instruments, distribution networks and art workers which are involved in the production and distribution of art works, and the assembly of their publics. Hence we might survey not just the formal properties of data projects or practices of prominent data workers (such as data scientists or data journalists), but a much broader cast of characters who are involved in the production, circulation and reception of data work.

Similar moves will be familiar from approaches inspired by Science and Technology Studies which view data infrastructures as *relations* of people, machines, software, standards, processes, practices, and cultures of knowledge production (e.g. Bowker and Star 1998, 2000; Star 1999; Star and Bowker 2002; Star and Ruhleder 1996; Jackson, Edwards, Bowker and Knobel 2007). Susan Leigh Star and Geoffrey Bowker suggest the notion of "infrastructural inversion" to bring neglected actors and processes into the foreground, including the role of non-human actors. In other recent work this has been framed in terms of "data assemblages" (see, e.g., Kitchin and Lauriault 2014).

One notable feature of many aspects of contemporary data politics is the emphasis on *redistributing* different forms of data work through digital technologies and networks. This redistribution comes in many different flavours. The tendency to redistribute "data work" from the public sector to the private sector is reflected in what Joseph Stiglitz calls the "default position" in information policy in the US, which is that states should not attempt what can be more effectively delivered by markets. This sentiment is also echoed in an influential paper called "Government Data and the Invisible Hand", suggesting that states cannot "keep pace" with the internet. This paper is picked up by Tim O'Reilly with his idea of "government as a platform" (which he opposed to "vending machine government"), an idea which was institutionalised as part of government policy in the UK (Gray 2014). Since the turn of the millennium, public information policy has seen an influx of different ideas concerning how and why to redistribute public data work – from enabling new kinds of innovation and businesses, reducing public sector costs, to crowdsourcing, distributed collaboration or peer production around data (modelled on open source software development), to grassroots, bottom up and participatory data cultures.

The redistribution of data worlds can be facilitated through a variety of devices and conventions, such as *open licenses* (like Creative Commons licenses); *data formats* such as Google Transit Feed Specification (later renamed General Transit Feed Specification); *online platforms* such as GitHub; *data portals* (such as data.gov); as well as hackathons, fellowships, and other public engagement activities. We may consider these not only as "transparency devices" (Barry 2010), but also as "infrastructuring devices" (Star and Bowker 2002; Pipek and Wulf 2009; Björgvinsson, Ehn and Hillgren 2010; Karasti 2014; Le Dantec and DiSalvo 2013), assembling different publics around data, whether it is to clean it up, crowdsource quality control of bus stop locations, monitor potholes, or make new apps and services. How these different forms of publicity, participation and contribution are materially organised is an important question which can be read in relation to recent research on the politics of openness and participation (Tkacz 2014) and of platforms, platformisation and platform capitalism (Helmond 2015; Srnicek 2016).

3. Data Worlds as Transnational Coordination

A third aspect of data worlds is world-making as transnational coordination, which includes projects of shaping, governing and articulating transnational relations, from empires and international institution building, to the networks, circuits and tendencies which are often studied by sociologists of globalisation (Sassen 2006).

Through this lens we can look at the world-making ambitions of legal and technical norms, standardisation, harmonisation and interoperability processes undertaken by a wide variety of different actors in the service of different projects for making things global. For example, UN bodies and EU statistical agencies have undertaken extensive programmes of work to align national forms of quantification – to support transnational policy coordination and comparison. Intergovernmental actors and international organisations such as the IMF, the World Bank and the

UN, have long supported the creation and alignment of systems and standards for the management of public finances in order to support objectives such as "fiscal discipline", the allocation and coordination of development funds, and the comparability of public spending across borders.

It is not only public institutions which share these kinds of world-shaping ambitions by means of data. They are accompanied by a host of researchers, companies, statisticians, consultants, analysts, accountants, scientists, activists, technologists, managers, journalists, ecologists, librarians, and others who seek to establish transnational information systems, practices, norms and standards. This may range from professional standards bodies (such as the International Accounting Standards Board), to multinational consultancies (such as Deloitte and other "big four" accounting firms), to private technology companies or startups (big tech companies such as Google to smaller projects like OpenCorporates), to non-profit and civil society initiatives (such as the Open Contracting Partnership's work on procurement data or Data2X's work on gender data).

Such initiatives often aim to shape the world through the coordination of data. Data worlds can make things amenable to measurement, monitoring, evaluation, analysis, and visualisation across space and time in support of diverse political, geopolitical, eco-political or political economic programmes – from neoliberal fiscal policy to market creation, gender equality to tax justice, increasing biodiversity to strengthening democracy. Civil society interventions to create and shape global data can be read in terms of other recent work around the history and sociology of quantification, as well as in terms of what some researchers have called "statactivism", and, more recently, "data activism" (Bruno, Didier and Vitale, 2014; Milan and van der Velden 2016). Longstanding information infrastructure projects, such as Amnesty International's "Urgent Actions" database, can be viewed as a kind of transnational "issue work", in order to render what might otherwise be disconnected incidents amenable to classification, measurement, comparison and virtual witnessing across borders.

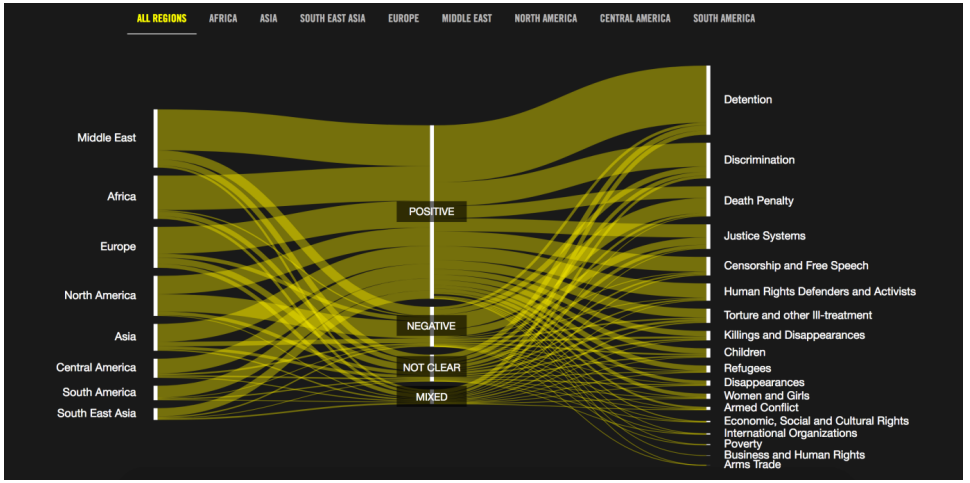

Figure 4. "Urgent Actions Visualised" from Amnesty International's Decoders project. http://decoders.amnesty.org/projects/decode-urgent-actions/results

There are of course many ways in which a given issue or matter of concern may be articulated, defined, parameterised, quantified, and given life through data. The transnational coordination aspect of data world-making is multivalent and may be considered in relation to a wide variety of utopian and dystopian, progressive and regressive political and ecological projects, as well as in terms of different histories and conceptions of land, territory, empire and earth. Here we may benefit from previous research on the colonial aspects of worlding in literary and cultural studies (e.g. Spivak 1985; Karagiannis and Wagner 2007; Clark, Finlay, and Kelly 2017); post-colonial computing (e.g. Irani, Vertesi, Dourish, Philip and Grinter, 2010); "planetary-scale computation" and emerging "technogeographies" (Bratton 2016; Gabrys 2016); as well as the surge of interest around "global intellectual history" and the making of worlds (e.g. Bell 2013).

We may consider data worlds to facilitate the demarcation and shaping of spaces, territories, environments, categories, identities and boundaries, separating interior from exterior, and sorting things, people and places out. They may also direct attention to different kinds of transnational issues, dynamics, concerns or collectives.

For example, we may look at the role of data worlds in relation to notions of the Anthropocene and the Capitalocene, in order to look at the role of human activity on a geological scale, as well as in the service of anthropologies of modernity, and projects to, as Bruno Latour puts it, "recompose a common world" (Kunkel 2017; Haraway 2016; Latour 2013). As well as deploying data worlds in order to better understand how human activity shapes the earth, information infrastructures may also be used to attempt to take various ecological signals into account in collectively redirecting its trajectory. As Goodman puts it: "if there is one world, it embraces a multiplicity of contrasting aspects; if there are many worlds, the collection of them all is one" (1978, 2). Data infrastructures can be used to establish the material character and limits of our one earth which contains such a plurality of social worlds and world-versions.

Data worlds can thus be understood as a means to institutionalise different forms of transnational coordination by providing the background against which things become seeable, sayable and doable with data across borders. Following recent research on neoliberal programmes (Roberts 2011; Davies 2014), data worlds may be considered as part of projects for reconfiguring relations between states, markets, citizens and civil society by foregrounding rankings, ratings and regimes of valuation in order to reinforce ideas of performance, competition and innovation, at the same time as moving tenets of economic and fiscal policy outside the realm of public and political deliberation. We can also read the redistributions of various forms of data work in terms of these contemporary imaginaries of democracy, markets and information – including those of competition, accountability, transparency, innovation, self-optimising systems and specific configurations of centralised management and decision-making coupled with decentralised delivery and contribution.

While there are indeed data worlds which may be configured to accelerate marketisation, bureaucratisation and what Habermas characterises as the "colonisation of the lifeworld"; other projects seek to address inequality and injustice, or to hold powerful elites accountable (as emphasised by the "statactivism" tradition), and all else in between. Data worlds can be malleable and may have unexpected consequences – such as in the cases of reports for investors being used by journalists and activists, or data from international development organisations being used by credit agencies.

Conclusion: Other Data Worlds Are Possible?

The aspects of data worlds described above are intended to gesture beyond two prominent narratives of data politics: of Promethean conceptions of liberating data as a resource on the one hand, and Orwellian visions of data surveillance, privacy and data protection on the other. These are vital parts of contemporary information politics, but there are other important aspects of what data is and what data does that should not be overlooked.

This article explores how theoretical traditions and literatures about worlds, worlding and world-making may be brought to bear to suggest different ways of thinking about data politics, highlighting three closely related aspects of data worlds. These three aspects are intended to be *illustrative* not exhaustive, and are intended as *overlapping* rather than distinct lenses through which to consider data infrastructures. They give rise to three different but closely related sets of questions for researching, theorising and reflecting on different aspects of data worlds.

1. **Data worlds as horizons of intelligibility**: What are the epistemic world-making capacities of data infrastructures? How might data infrastructures be involved in "making things up"? Can they provide conditions of possibility for different ways of seeing, saying and knowing collective life, and if so, how?
2. **Data worlds as collective accomplishments**: Who and what is involved with making, and making sense with, data? How are data worlds being redistributed through digital technologies? Who is (and who isn't) able to shape data worlds? What kinds of practices of participation and public involvement are emerging around data worlds?

3. **Data worlds as transnational coordination**: How might data infrastructures be implicated in different attempts to "make things global"? What kinds of transnational alliances and circuits are being formed and to what end? Who advocates which kinds of data worlds, and according to which kinds of visions and fields of transnational coordination (from international relations to earth science)?

It is worth noting that it remains an empirical question as to *how* and *to what extent* data infrastructures are involved in world-making in these three senses. Data infrastructures can be deployed with certain epistemic, social and political aspirations and imaginaries in mind which they do not live up to. Data projects can fail to become data worlds in these three senses.

The notion of data worlds is not just intended to advance research on data politics. Following recent debates about the performativity of critique (e.g. Latour 2004), and calls to integrate critical, theoretical and humanistic reflection into technical practice (e.g. Agre 1997; Rieder and Röhle 2012; Berry 2014), I am particularly interested in how the notion of data worlds might suggest different kinds of data politics. Of course, theory and critique can contribute to doing things differently, as critical data studies researchers have pointed out. Dalton, Taylor and Thatcher, for example, propose to "develop alternative knowledges that reflect and build on our criticisms" (2016).

To this end, I'd like to propose the notion of "critical data practice" as a site for pedagogical experimentation, research and intervention around the politics of data. This follows Agre's notion of "critical technical practice" which he uses to characterise his attempts to integrate historical and theoretical reflection around artificial intelligence into his work as an AI researcher (Agre 1997). The crucial question here is *what difference* critical studies can make in doing things with data. As well as contributing to critical genealogies and sociologies of the politics of "data worlds" and "data world-making" projects, researchers and universities might contribute to "making space" for such experimentation and intervention around public data infrastructures and the role they play in collective life.

The three aspects of data worlds I have examined are intended to assist with the task of rethinking the politics of public data, by considering *how* and *for whom* it is made public. Thus we may examine the organisation of what Evelyn Ruppert calls "data publics" (Ruppert 2015) beyond a focus on accessing, liberating and using data, and taking a broader look at how different actors engage with, mobilise around, shape and are shaped by, public data infrastructures. This includes distributed collaboration around different kinds of "data work" – from projects inspired by free software, free culture and open access movements such as Open Street Map or Wikidata, to data journalism and data activism projects for counting police killings or migrant deaths, to other kinds of civil society interventions for changing the socio-technical arrangements by which public institutions account for issues by means of data. As well as attending to these arrangements, researchers may also consider "experiments in participation" (Lezaun, Marres and Tironi 2016; Marres 2012) around data worlds, which are also cognisant of patterning and politics of these participatory processes. Such experimentation would not just aim to interpret data worlds, but also to question them, to re-imagine them, and to change them.

Acknowledgements

I'd like to thank Frank Pasquale, Will Davies, Liz McFall, Daniel Wilson and other participants at a workshop on "Outnumbered! Statistics, Data and the Public Interest" at the Centre for Research in the Arts, Social Sciences and Humanities (CRASSH), University of Cambridge in June 2017 for their useful feedback on an earlier version of this paper. It has also benefitted from feedback and input from colleagues in talks and workshops at King's College London, the University of Amsterdam, the Politecnico di Milano, the University of Westminster, the University of Siegen and the Université Paris Nanterre. I'm also most grateful for time, suggestions and encouragement from Claudia Aradau, Liliana Bounegru, Carolin Gerlitz, Noorje Marres, John Durham Peters and two anonymous reviewers. Parts of the research for this article were made possible by a Starting Grant of the European Research Council (639379-DATACTIVE, https://data-activism.net) at the University of Amsterdam, where I was a postdoctoral researcher (2015-2016) and continue as a Research Associate; as well as through a Prize Fellowship at the Institute for Policy Research, University of Bath (2016-2017).

References

Agre, Philip E. 1997. "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI." In *Bridging the Great Divide: Social Science, Technical Systems, and Cooperative Work*, edited by Geoffrey Bowker, Les Gasser, Susan Leigh Star, and Bill Turner, 131-157. Mahwah, NJ: Erlbaum.

Apel, Karl-Otto. 1973. *Transformation der Philosophie*, Volume 1. Frankfurt: Suhrkamp.

Barry, Andrew. 2010. "Transparency as a Political Device." In *Débordements: Mélanges offerts à Michel Callon*, edited by Madeleine Akrich, Yannick Barthe, Fabian Muniesa, and Philippe Mustar, 21-40. Paris: Presses des Mines.

Becker, Howard. 1984. *Art Worlds*. Berkeley, CA: University of California Press.

Bell, Duncan. 2013. "Making and Taking Worlds." In *Global Intellectual History* edited by Samuel Moyn and Andrew Sartori, 254-279. New York: Columbia University Press.

Benjamin, Walter. 1996. *Walter Benjamin: Selected Writings 1913 - 1926* (Vol. 1). Cambridge, MA: Harvard University Press.

Benjamin, Walter. 1999. *The Arcades Project*. Edited by Rolf Tiedemann. Translated by Howard Eiland and Kevin McLaughlin. Cambridge, MA: Harvard University Press.

Berry, David M. 2014. *Critical Theory and the Digital*. London: Bloomsbury.

Björgvinsson, Erling, Pelle Ehn, and Per-Anders Hillgren. 2010. "Participatory Design and "Democratizing Innovation"." In *Proceedings of the 11th Biennial Participatory Design Conference*, 41–50. PDC '10. New York: ACM. https://doi.org/10.1145/1900441.1900448.

Bloor, David. 2002. *Wittgenstein, Rules and Institutions*. London: Routledge.

Bloor, David. 1983. *Wittgenstein: A Social Theory of Knowledge*. London: Macmillan.

Blok, Anders, and Morten Axel Pedersen. 2014. "Complementary Social Science? Quali-Quantitative Experiments in a Big Data World." *Big Data & Society*, 1(2).

Bowker, Geoffrey C. 2014. "The Infrastructural Imagination." In *Information Infrastructure(s): Boundaries, Ecologies, Multiplicity* edited by Alessandro Mongili and Giuseppina Pellegrino, xii-xiii. Cambridge, UK: Cambridge Scholars Publishing.

Bowker, Geoffrey C., and Susan Leigh Star. 1998. "Building Information Infrastructures for Social Worlds – The Role of Classifications and Standards." In *Community Computing and Support Systems* edited by Toru Ishida, 231-248. Berlin: Springer.

Bowker, Geoffrey C., and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

Bowie, Andrew. 2013. *Aesthetics and Subjectivity: From Kant to Nietzsche*. Manchester: Manchester University Press.

Boyer, Dominic. 2013. *The Life Informatic: Newsmaking in the Digital Era*. Ithaca: Cornell University Press.

Bratton, Benjamin H. 2016. *The Stack: On Software and Sovereignty*. Cambridge, MA: MIT Press.

Bruno, Isabelle, Florence Jany-Catrice, and Beatrice Touchelay, eds. 2016. *The Social Sciences of Quantification: From Politics of Large Numbers to Target-Driven Policies*. New York: Springer.

Clark, Tom, Emily Finlay, and Philippa Kelly, eds. 2017. *Worldmaking: Literature, Language, Culture*. Amsterdam: John Benjamins Publishing Company.

Clarke, Adele E., and Susan Leigh Star. 2008. "The Social Worlds Framework: A Theory/Methods Package." In *The Handbook of Science and Technology Studies*, edited by Edward J. Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman, 113–38. Cambridge, MA: MIT Press.

Dalton, Craig M., Linnet Taylor, and Jim Thatcher. 2016. "Critical Data Studies: A Dialog on Data and Space." *Big Data & Society* 3 (1).

Davies, Will. 2014. *The Limits of Neoliberalism: Authority, Sovereignty and the Logic of Competition*. London: SAGE Publications.

Desrosières, Alain. 2002. *The Politics of Large Numbers: A History of Statistical Reasoning*. Translated by Camille Naish. Cambridge, MA: Harvard University Press.

Espeland, Wendy Nelson, and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology / Archives Européennes de Sociologie* 49 (3): 401–436.

Foucault, Michel. 1972. *Archaeology of Knowledge*. Translated by A. M. Sheridan Smith. New York: Pantheon Books.

Flyverbom, Mikkel, Anders Koed Madsen, and Andreas Rasche. 2017. "Big Data as Governmentality in International Development: Digital Traces, Algorithms, and Altered Visibilities." *The Information Society* 33 (1): 35–42.

Gabrys, Jennifer. 2016. *Program Earth*. Minneapolis: University of Minnesota Press.

Geuss, Raymond. 2003. *Public Goods, Private Goods*. Princeton, NJ: Princeton University Press.

Giaccardi, Elisa, Nazli Cila, Chris Speed, and Melissa Caldwell. 2016. "Thing Ethnography: Doing Design Research with Non-Humans." In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 377–387. DIS '16. New York, NY, USA: ACM. https://doi.org/10.1145/2901790.2901905.

Gillespie, Tarleton. 2010. "The Politics of 'Platforms'." *New Media & Society*, 12(3): 347–364.

Goodman, Nelson. 1978. *Ways of Worldmaking*. Indianapolis, IN: Hackett Publishing.

Gray, Jonathan. 2014. "Towards a Genealogy of Open Data." General Conference of the European Consortium for Political Research. http://dx.doi.org/10.2139/ssrn.2605828

Gray, Jonathan. 2016. "Datafication and Democracy: Recalibrating Digital Information Systems to Address Societal Interests." *Juncture*, 23(3). Retrieved from http://www.ippr.org/juncture/datafication-and-democracy

Gray, Jonathan. 2018. "Quand les mondes de données sont redistribués: Open Data, infrastructures de données et démocratie." *Statistique et Société*, 5(3).

Gray, Jonathan, Danny Lämmerhirt, and Liliana Bounegru. 2016. "Changing What Counts: How Can Citizen-Generated and Civil Society Data Be Used as an Advocacy Tool to Change Official Data Collection?" CIVICUS and Open Knowledge. http://papers.ssrn.com/abstract=2742871.

Hacking, Ian. 1984. "Wittgenstein Rules." *Social Studies of Science*, 14 (3).

Hacking, Ian. 1985. "Making People Up." In *Reconstructing Individualism: Autonomy, Individuality and the Self in Western Thought*, edited by Thomas C. Heller, Morton Sosna and David E. Wellbery, 222-236. Stanford, CA: Stanford University Press.

Hacking, Ian. 1990. *The Taming of Chance*. Cambridge, UK: Cambridge University Press.

Hacking, Ian. 2002. "How, Why, When and Where Did Language Go Public?" In *Historical Ontology*, 121-139. Cambridge, MA: Harvard University Press.

Halpern, Orit, Jesse LeCavalier, Nerea Calvillo, and Wolfgang Pietsch. 2013. "Test-Bed Urbanism." *Public Culture* 25 (2): 272–306.

Halpern, Orit. 2015. *Beautiful Data: A History of Vision and Reason Since 1945*. Durham: Duke University Press Books.

Haraway, Donna J. 2016. *Staying with the Trouble: Making Kin in the Chthulucene*. Durham: Duke University Press Books.

Helmond, Anne. 2015. "The Platformization of the Web: Making Web Data Platform Ready." *Social Media + Society*, 1(2).

Iliadis, Andrew, and Federica Russo. 2016. "Critical Data Studies: An Introduction." *Big Data & Society* 3 (2).

Irani, Lilly, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E. Grinter. 2010. "Postcolonial Computing: A Lens on Design and Development." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1311–1320. CHI '10. New York, NY, USA: ACM.

Jackson, Steven J., Paul N. Edwards, Geoffrey C. Bowker, and Cory P. Knobel. 2007. "Understanding Infrastructure: History, Heuristics and Cyberinfrastructure Policy." *First Monday* 12 (6).

Karagiannis, Nathalie, and Peter Wagner, eds. 2007. *Varieties of World-Making: Beyond Globalization*. Liverpool: Liverpool University Press.

Karasti, Helena. 2014. "Infrastructuring in Participatory Design." In *Proceedings of the 13th Participatory Design Conference: Research Papers-Volume 1*, 141–150. ACM.

Kelty, Chris M. 2008. *Two Bits: The Cultural Significance of Free Software: The Cultural Significance of Free Software and the Internet*. Durham, NC: Duke University Press.

Kitchin, Rob, and Tracey P. Lauriault. 2014. "Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work." SSRN Scholarly Paper ID 2474112. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2474112.

Kunkel, Benjamin. 2017. "The Capitalocene." *London Review of Books*, 2 March 2017.

Lampland, Martha, and Susan Leigh Star. 2009. *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.

Latour, Bruno. 2013. *An Inquiry into Modes of Existence: An Anthropology of the Moderns*. Translated by Catherine Porter. Cambridge, Massachusetts: Harvard University Press.

Latour, Bruno. 2004. "Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern." *Critical Inquiry*, 30 (2).

Lavigne, Sam. 2017. "Taxonomy of Humans According to Twitter." *The New Inquiry*. 5 July 2017. https://thenewinquiry.com/taxonomy-of-humans-according-to-twitter/.

Law, John. 2009. "Seeing Like a Survey." *Cultural Sociology*, 3 (2).

Law, John, Evelyn Ruppert, and Mike Savage. 2011. "The Double Social Life of Methods." Other 95. CRESC Working Paper. Milton Keynes: Open University.

Dantec, Christopher A Le, and Carl DiSalvo. 2013. "Infrastructuring and the Formation of Publics in Participatory Design." *Social Studies of Science* 43 (2): 241–64.

Lezaun, Javier, Noortje Marres, and Manuel Tironi. 2016. "Experiments in Participation." In *Handbook of Science and Technology Studies: Fourth Edition*, edited by Ulrike Felt, Rayvon Fouche, Clark A. Miller, and Laurel Smitt-Doer, Fourth Edition, 195-221. Cambridge: MIT Press.

Lifschitz, Avi. 2012. *Language and Enlightenment: The Berlin Debates of the Eighteenth Century*. Oxford: Oxford University Press.

Lynch, Michael. 1992. "Extending Wittgenstein: The Pivotal Move from Epistemology to the Sociology of Science." In *Science as Practice and Culture*, edited by Andrew Pickering, 2nd Edition, 215-265. Chicago, IL: University of Chicago Press.

Mackenzie, Donald. 2008. *An Engine, Not a Camera: How Financial Models Shape Markets*. Cambridge, MA: MIT Press.

Marres, Noortje. 2012. *Material Participation: Technology, the Environment and Everyday Publics*. London: Palgrave Macmillan.

Milan, Stefania and Lonneke van der Velden. 2016. "The Alternative Epistemologies of Data Activism." *Digital Culture & Society*, 2(2).

Mol, Annemarie. 1999. "Ontological Politics. a Word and Some Questions." *The Sociological Review* 47 (S1): 74–89.

Pipek, Volkmar and Volker Wulf. 2009. "Infrastructuring: Toward an Integrated Perspective on the Design and Use of Information Technology." *Journal of the Association for Information Systems* 10 (5).

Porter, Theodore M. 1986. *The Rise of Statistical Thinking, 1820-1900*. Princeton, NJ: Princeton University Press.

Porter, Theodore M. 1996. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.

Rieder, Bernhard, and Theo Röhle. 2012. "Digital Methods: Five Challenges." In *Understanding Digital Humanities*, edited by David M. Berry, 67–84. London: Palgrave Macmillan.

Roberts, Alasdair. 2011. *The Logic of Discipline: Global Capitalism and the Architecture of Government*. Oxford: Oxford University Press.

Rottenburg, Richard, Sally E. Merry, Sung-Joon Park, and Johanna Mugler, eds. 2015. *The World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge, UK: Cambridge University Press.

Ruppert, Evelyn. 2015. "Doing the Transparent State." In *The World of Indicators: The Making of Governmental Knowledge through Quantification*, edited by Richard Rottenburg, Sally E. Merry, Sung-Joon Park and Johanna Mugler, 127-150. Cambridge, UK: Cambridge University Press.

Ruppert, Evelyn, John Law, and Mike Savage, eds. 2013. ""The Social Life of Methods", Special Issue'." *Theory, Culture & Society* 30 (4).

Sassen, Saskia. 2006. *A Sociology of Globalization*. New York: W.W. Norton.

Scott, James. 1999. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.

Singer, Brooke. 2016. "A Chronology of Tactics: Art Tackles Big Data and the Environment." *Big Data & Society*, 3(2).

Spivak, Gayatri Chakravorty. 1985. "Three Women's Texts and a Critique of Imperialism." *Critical Inquiry* 12 (1): 243–61.

Srnicek, Nick. 2016. *Platform Capitalism*. London: Polity Press.

Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–91.

Star, Susan Leigh and Geoffrey Bowker. 2002. "How to Infrastructure." In *Handbook of New Media: Social Shaping and Consequences of ICTs*, edited by Leah A. Lievrouw and Sonia Livingstone, 151-162. London: Sage.

Star, Susan Leigh, and Karen Ruhleder. 1996. "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces." *Information Systems Research* 7 (1): 111–34.

Star, Susan Leigh, Geoffrey Bowker and Laura J. Neumann. 1997. "Transparency at Different Levels of Scale: Convergence between Information Artifacts and Social Worlds." Available at: https://www.ics.uci.edu/~gbowker/converge.html

Strauss, Anselm. 1978. "A Social World Perspective." *Studies in Symbolic Interaction* 1.

Taylor, Charles. 1985. *Philosophical Papers: Volume 1, Human Agency and Language*. Cambridge, UK: Cambridge University Press.

Taylor, Charles. 2016. *The Language Animal: The Full Shape of the Human Linguistic Capacity*. Cambridge, MA: Harvard University Press.

Tkacz, Nathaniel. 2014. *Wikipedia and the Politics of Openness*. Chicago, IL: University Of Chicago Press.

Tuschling, Anna. 2016. "Historical, Technological and Medial a Priori: On the Belatedness of Media." *Cultural Studies* 30 (4): 680–703.

Verran, Helen. 2015. "Enumerated Entities in Public Policy and Governance." In *Mathematics, Substance and Surmise*, edited by Ernest Davis and Philip J. Davis, 365–79. New York: Springer, Cham.

Winthrop-Young, Geoffrey, Ilinca Iurascu, and Jussi Parikka, eds. 2013. ""Cultural Techniques". Special Issue." *Theory, Culture & Society* 30 (6). http://journals.sagepub.com/toc/tcsa/30/6.

Wittgenstein, Ludwig. 2001. *Philosophical Investigations: The German Text, with a Revised English Translation*. Oxford: Blackwell.

Biography

Jonathan Gray
Dr. Jonathan Gray is Lecturer in Critical Infrastructure Studies at the Department of Digital Humanities, King's College London, where he is currently writing a book on data worlds and the politics of public information. He is also Cofounder of the Public Data Lab; and Research Associate at the Digital Methods Initiative (University of Amsterdam) and the médialab (Sciences Po, Paris). More about his work can be found at jonathangray.org and he tweets at @jwyg.